A recurrent multilayer model with Hebbian learning and intrinsic plasticity leads to invariant object recognition and biologically plausible receptive fields

Michael Teichmann, Fred H. Hamker

Much effort has been spent to develop biologically plausible models for different aspects of the processing in the visual cortex. However, most of these models are not investigated with respect to the functionality of the neural code for the purpose of object recognition comparable to the framework of deep learning in the machine learning community.

We developed a model of V1 and V2 based on anatomical evidence of the layered architecture (Douglas & Martin, 2004, Potjans & Diesmann, 2012), using excitatory and inhibitory neurons where the connectivity to each neuron is learned in parallel. For V2 we use the same inner area connectivity as for V1. We address learning by three different mechanisms of plasticity: intrinsic plasticity, Hebbian learning with homeostatic regulations, and structural plasticity.

Intrinsic plasticity solves the problem of imbalanced encoding of the neurons by controlling the amount each single neuron participates in the encoding of stimuli. It allows us to build deep networks while preserving the coding quality. It adopts the excitability of a neuron through an adaptive activation threshold and slope (Triesch, 2005). The threshold change is based on the relation of the neuron's activity to the average layer activity. The slope change considers the relation of the neuron's squared activity to the average of the squared activity of all neurons within a layer.

The tuning of receptive fields is specified by Hebbian learning (Teichmann et al.,2012; Wiltschut & Hamker, 2009). In order to learn invariance properties, our learning rule uses the postsynaptic calcium concentration rather than the neuronal activity. As the calcium concentration reflects an activity trace, the temporal proximity of inputs maps to the selectivity of postsynaptic neurons. The length of this trace is short for layer-4 neurons (typically simple cells) and sufficiently long for layer-2/3 neurons (typically complex cells). To normalize the weight vectors we use a homeostatic regulation (Oja, 1982) dynamically adjusting the length of the weight vector based on the activity range of the respective neuron.

The problem that the connection structure, and thus the receptive field size, can largely differ between individual cortical neurons is solved in our model by an activity dependent form of structural plasticity. The mechanism exploits the spatial organization of the neurons to allow that the connectivity can grow or retract to an individually amount for each neuron. This mechanism reduces the need of a priori knowledge about the amount of the neuron's afferents, and thus reduces the modelers bias for its receptive field development.

For evaluation, we trained the model on natural scenes and analyzed the resulting receptive fields, i.e. invariance properties, tuning characteristics and robustness to input variations, being of essential interest as they are showing the basic principles of feature extraction and invariant representation for object recognition. Further, we show the recognition accuracies of the model trained on natural scenes to the COIL-100 dataset (COIL-100) for all V1 and V2 layers to demonstrate the object recognition capability of each model layer.

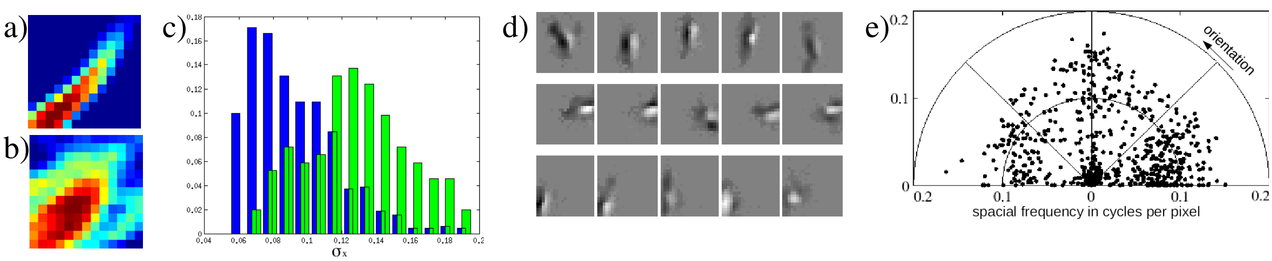

Figure 1: Receptive field evaluation examples. Spacial tuning to the optimal stimulus for a (a) simple cell, (b) complex cell. (c) Histogram of the tuning widths for simple (blue) and complex cells (green). (d) Receptive fields of three complex cells (row), illustrated by the receptive fields of the five most strongly connected simple cells. (e) Orientation and frequency selectivity of simple cells.